

**Original citation:**

Tillson, John. (2017) The problem of rational moral enlistment. Theory and Research in Education.

**Permanent WRAP URL:**

<http://wrap.warwick.ac.uk/88567>

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

This had been posted ahead of publication.

Tillson, John. (2017) The problem of rational moral enlistment. Theory and Research in Education. Copyright © 2017 The Authors. Reprinted by permission of SAGE Publications.  
<https://doi.org/10.1177/1477878517716427>

**A note on versions:**

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP url' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk)

# THE PROBLEM OF RATIONAL MORAL ENLISTMENT

J. Tillson

University of Warwick, UK

## ABSTRACT

How can one bring children to recognize the requirements of morality without resorting only to non-rational means of persuasion (i.e. what rational ground can be offered to children for their moral enlistment)? Michael Hand has recently defended a foundationalist approach to answering this question and J. White has responded by a) criticizing Hand's solution to the Problem of Rational Moral Enlistment, and b) attempting to circumvent the problem by suggesting a Humean route which understands moral enlistment as grounded in sentiment. While I do not accept Hand's preferred solution to the Problem of Rational Moral Enlistment, I am also unpersuaded by White's attempt to circumvent it. Instead, making use of work by Ben Spiecker and Jan Steutel (2001), I attempt a different solution to the Problem of Rational Moral Enlistment – one appealing to reflective equilibrium rather than to ethical foundationalism as Hand's does. Whereas Hand hopes to ground rational moral enlistment in a single, self-evident foundational justification of some moral standards, I instead hope to facilitate rational moral enlistment through a rational procedure which starts with students' existing moral commitments and attempts to revise or expand them through a certain kind of critical reflection.

## KEYWORDS

Moral education; indoctrination; foundationalism; reflective equilibrium; rationality; persuasion.

## 1. INTRODUCTION

In his inaugural professorial lecture published under the title 'Towards a Theory of Moral Education,' Michael Hand reformulated an old problem of moral education, and proposed a new solution to it (2014b). The problem is this: how can one educate children to recognize the requirements of morality without resorting to non-rational means of persuasion alone (i.e. what rational ground can be offered to

children for their moral enlistment)? Call this the Problem of Rational Moral Enlistment. His answer is that *some* moral standards ‘are robustly justified,’ finding their grounds for rational persuasion in the amelioration of ‘breakdowns in cooperation and outbreaks of conflict’ within social groups (2014b: 528, 530). Only these justifications can be readily understood by students, and so form the basis of a rational initiation into recognizing the requirements of morality, or ‘rational moral enlistment’, as I shall call it. According to Hand, these are the only moral requirements that can be robustly justified. J. White recently a) criticized Hand’s solution to the Problem of Rational Moral Enlistment, and b) attempted to circumvent the problem by suggesting a Humean route in which moral motivation is properly grounded in sentiment, with moral enlistment properly done by modelling and praising the desired sentiments (White, 2016: 454). While I do not accept Hand’s preferred solution to the Problem of Rational Moral Enlistment, I am also unpersuaded by White’s attempt to circumvent the problem. Part of my disagreement with White is motivated by a broadly Handian conception of education.

In the next Section of this paper, I explain why educators should prefer to effect moral enlistment with rational justifications. In Section 3, I further explain Hand’s proposed solution to the Problem of Rational Moral Enlistment (i.e. the robust justification he claims to have found for some moral standards which may be used in students’ rational moral enlistment). Section 4 draws on White’s critique, arguing that Hand does not manage to provide a self-evident justification to ground rational moral enlistment. In Section 5 I resist White’s attempts to circumvent the Problem of Rational Moral Enlistment. Section 6, drawing on work by Ben Spiecker and Jan Steutel (2001), describes the process of reflective equilibrium. In Section 7, I attempt a different solution to the Problem of Rational Moral Enlistment – one appealing to reflective equilibrium rather than to ethical foundationalism as Hand’s does. Finally, in Section 8, I respond to some possible objections. In brief, whereas Hand hopes to ground rational moral enlistment in a single, self-evident foundational justification of some moral standards, I instead hope to ground rational moral enlistment in a rational procedure which starts with students’ existing moral commitments and attempts to revise or expand them through a certain kind critical reflection. This procedure consists, in short, in articulating and collecting those ethical judgments to which the student is most stably committed, developing a best fit between particular judgements and general

principles that would explain them, and a best fit between these judgements and principles on the one hand, and wider background knowledge on the other.

## 1. THE EPISTEMIC CRITERION AND RATIONAL MORAL ENLISTMENT

Hand understands the Problem of Rational Moral Enlistment in a way that is systematically integrated with his views on education and indoctrination quite generally, and this will require a little scene setting at the outset. Hand, draws a fundamental distinction between directive teaching and non-directive teaching. With directive teaching, one intends to promote particular cognitive attitudes to some proposition (whether belief, disbelief, or agnosticism). With non-directive teaching, one does not intend to promote any particular cognitive attitudes to the proposition introduced beyond simply understanding it (Hand, 2014a). Hand describes two limitations on what may be directive taught, one regarding content, and one regarding methods, together these constitute the 'Epistemic Criterion'. Regarding content: teachers are entitled to impart belief of a proposition if and only if disbelief of and agnosticism towards that proposition are 'contrary to reason' given the existing state of evidence and argument, i.e. where the proposition enjoys decisive probative support. Examples of such propositions include that smoking causes cancer, using sunscreen prevents skin cancer, anthropogenic climate change is happening, and that evolution by natural selection is responsible for the proliferation of life on earth. Regarding method: the directive teaching is by means of rational persuasion, such that the reasons for the belief can be understood by the pupils. According to Hand, beliefs can be transmitted in two rational ways. First, 'where beliefs are known to be true, they can be imparted by means of rational demonstration' in which a proposition is proven to enjoy sufficient probative force to make non-belief irrational (Hand, 2003: 545). Second, 'other things being equal, when a person perceived by others to be an intellectual authority [on a relevant matter] asserts that a proposition is true, she places them under a rational obligation to accept her assertion. [Insofar as they are rational] she *imparts a belief* to her listeners, and she does so by *appealing to their reason*' (Hand, 2003: 551).

It must be stressed however, that obeying the Epistemic Criterion is not a strict moral requirement in that it can sometimes be trumped by other moral considerations. Indeed, where Rational Moral Enlistment is impossible, provided moral knowledge is had, Moral Indoctrination might be obligatory.

Consider the legion crimes which continue to be perpetrated in the name of the Islamic State of Iraq and the Levant (ISIL), and its Salafi form of jihadism. These are widely known, from the beheadings of aid workers, police officers, soldiers, civilians, journalists and brutal forms of execution for adulterers and homosexuals, often filmed and distributed, to their “systematic destruction of ancient cultural and archaeological heritage sites and artefacts” (Muir, 2016). So as to end further such atrocities, if any security force had the means to convert ISIL’s members to reject its extremist worldview by using non-rational methods, they would be morally obliged to use them in this way. It might be desirable to supplement rational methods of moral enlistment with non-rational methods of moral enlistment (such as praise and blame, reward and punishment, and perhaps even shaming) and it might equally be desirable to guide behaviour to be within the limits of morality with such non-rational methods (where cognitive acquiescence to these requirements is not the aim).

What can be said to justify the constraints of content and method that Hand places on what may be directly taught? Let’s take the rational persuasion constraint first. Propositions (moral propositions included) ought to be imparted by rational means for two central reasons, but there is a third additional reason for imparting moral beliefs by rational means as we shall see. First, those propositions which may be regarded as contrary to reason to deny will vary as our information base changes, and if it is better for us to be right about these matters, it is important that one’s propositions update with that information base. This requires us to believe for relevant reasons, and update our belief set as those reasons dematerialize or further reasons materialize (even if this is simply a shift in expert consensus). We cannot then simply be given a set of irrationally held beliefs that can reliably shift to fit the truth. Second, one cannot simply hive off areas of curriculum content to be taught rationally and others to be taught non-rationally, since, as Jonathan Lowe puts it, ‘Truth is single and indivisible or, to put it another way, the world or reality as a whole is unitary and necessarily self-consistent’ (Lowe, 2002: 3). As all liars know, lies snowball with time in order to maintain the original lie, and the case of false beliefs is just like this. Bad reasons are wont to imply and often generate further false beliefs. Finally, in the case of moral beliefs in particular, there is the further problem of not deserving the moral credit if one does the right thing dumbly.

Hand's constraint on method seems to be highly defensible, then. What can be said to justify Hand's constraint on content for deciding what may be directly taught? Some might hope to teach some false beliefs for the point of view of generating utility. For instance, Plato's noble lie was designed to foster a sense of social solidarity. However, again one cannot hive off areas of curriculum content to impart falsehoods and others impart truths since, again 'Truth is single and indivisible,' false beliefs will imply and generate further false beliefs. One cannot hope to cultivate false beliefs by non-rational means in an isolated, restricted domain, the falsehoods will bleed out into other domains, and other truths will be hostage to maintenance of some presumed isolated falsehood. While one might bring children to endorse a view for non-rational reasons at first, with the hope that they will come to see its rational basis in due course, this too easily risks the proliferation of further views on the basis of non-rational reasons. Furthermore, one cannot hope to simply backtrack, since it is not as if when one discovers what the first lie was, all the others that were contingent on it would then disappear as with an instantaneously error correcting computer program. So there will be collateral damage caused by any such undertaking.

Others might regard Hand's requirements as too restrictive for another reason. While they might agree with Hand that we ought not to cultivate assent to propositions that it is contrary to reason to believe, they will maintain that we ought to cultivate assent to some propositions that can reasonably be denied. For instance, one might argue that we ought to rationally persuade children to believe that they ought, for the sake of future generations, not to damage the natural environment, even if there is no knock down argument against those who claim that it is impossible to wrong future generations. While there may be no knock down argument against such a view, denying it may be very much more plausible than maintaining it, and this marked difference in plausibility can justify promoting the view as markedly more plausible. However, where there is no such pronounced difference in plausibility, the following consideration seems apposite: suppose one has two rival theories, both of which are equally plausible, but one of which if not acted on would lead to a worse result, it seems that while one has no epistemic reason to believe the one over the other, one has reason to act on one rather than the other, and to encourage others to act on theory rather than the other. Supposing that epistemically there were a 50/50 chance as to whether we have any obligations to future generations, if the burdens of acting as

if we had obligations to future generations were smaller than the cost in suffering to future generations of failing to act in this way, we could sensibly persuade children to act as if they had these obligations, without attempting to persuade them that they do have these obligations.

Hand believes that for a subsection of moral beliefs, we can provide decisive ground for saying that they are true, for another subsection, we can provide grounds for saying that they are false, and for more still, enough of a case can be made in their favour that they should be included on the curriculum and taught in a non-directive fashion. It is the supposedly decisive grounds that Hand offers that are the subject of our next section.

We should note before continuing that in framing the problem, it seems that Hand has a prior commitment to averting indoctrination, and to facilitating education, and we may want to ask whether these are morally required of teachers. I would suggest that it is a moral wrong to indoctrinate, and that it would seem strange to on the one hand admit to scepticism about moral standards, and on the other be committed to a task like educating and losing sleep about whether one is indoctrinating moral standards. Without moral knowledge, the problem of indoctrination loses its sting. However, since Hand thinks that there is genuine moral knowledge to be had (that some reasons genuinely do justify some moral standards), this concern with indoctrination can be vindicated.

## 2. HAND'S FOUNDATIONAL REASON

Notice that all of the foregoing discussion about directive and non-directive teaching has concerned forming beliefs; not habits or dispositions.<sup>1</sup> For Hand, any moral education is incomplete which fails to teach for cognitive assent to such moral judgements as 'it is wrong to kill just for the sake of it':

If, however, the aim of moral education is to bring about full moral commitment ... Children must be brought both to subscribe to moral standards and to believe their subscription to be justified. It will not be enough for them to intend not to steal, expect others not to steal, and know how to avoid stealing. They will also need to believe that at least one putative justification for the moral prohibition on stealing is sound (Hand, 2014b: 527)

All it remains for us to do is see whether we have the means to rationally convince children that morality properly makes certain demands of them, and to submit to those demands. Hand, takes a foundationalist approach to moral epistemology. According to such accounts, ‘a moral belief is justified if and only if it is either self-justifying or bears an appropriate inferential relation to a belief that is self-justifying’ (McMahan, 2000: 110). To fill this role of a self-justifying moral belief, Hand appeals to what he calls ‘the problem of sociality’. The basic problem is this: if we all did whatever we wanted and our wants were not mutually beneficial, we would most of us, mostly have a miserable time – even the strongest, smartest and most powerful would not have good prospects for doing well.

Since all of us who live in social groups have an interest in averting breakdowns of cooperation and outbreaks of conflict, so we all have an interest in holding ourselves and others to rules of conduct that sustain cooperation and peace (ibid: 529)

### 3. RESISTING HAND’S FOUNDATIONALISM

One might worry that the moral foundation that Hand finds in ameliorating the problem of sociality is 1) less persuasive than the prohibitions, obligations and other standards that it is meant to ground, 2) fails to ground intuitions that are more appealing than this way of grounding them, and 3) conflicts with other intuitions that are more appealing than this way of grounding them. To take these points in order, first, it is easier to agree that I have reason not murder others than it is to agree that the reason I have not to murder others is the amelioration of the problem of sociality. Second, it seems very plausible that we ought not to torture animals, but the amelioration of the problem of sociality offers us no reason to think so. Again, it seems that we would have reason (if not decisive reason) to refrain from killing an innocent person even if doing so would somehow avert social breakdown: a reason that *ex hypothesi* cannot be grounded in evading the possibility of social breakdown. Third, it seems immoral to farm and slaughter non-human animals in order to eat them even if doing so slightly ameliorated the problem of sociality.

What the problem of sociality motivates via a concern with one’s own self-interest is a social contract, which cannot be identified with characteristically moral reasons. Firstly, social contracts



presume morality to some extent (notably that one ought to honour one's contractual commitments), and so cannot completely ground it. Secondly, while self-interest overlaps with moral concerns in that one's wellbeing is of moral concern, we can have reason in the shape of the wellbeing of others to promote their concerns where that is not always to be defended in a practical advantage to myself. Hand's strategy explains why it can be mutually beneficial to show mutual regard, but not why there are demands of morality that go beyond self-regard, and indeed noble deeds which go beyond self-regard (which may be supererogatory – such as self-sacrifice). Indeed, as White observes, if self-interest is intended to motivate mutual cooperation, or ground the requirements of morality, then Hand's theory fails to address the problem of the free rider (White, 2016: 453). That is, it fails to address the sneaky strategy one might have of (in one's self interest) disregarding the requirements of morality where it will go unnoticed, while benefitting from other people following them. The particular form of the problem here is not that this strategy is merely possible, and hard to dissuade others from, but that there does not seem to be anything wrong with the strategy if self-advantage is the ultimate bedrock of morality.

When I deny that Hand has *identified* a foundational moral truth, I do not deny that there *are* any foundational moral truths. Not knowing of any self-evident moral truths myself, what motivates my turn to reflective equilibrium is the prospect that no foundational moral truths are *required* in order to make rational moral enlistment possible. Before advocating this view, I want to discuss J. White's attempt to circumvent the Problem of Rational Moral Enlistment.

#### 4. WHITE AND THE NON-PROBLEM OF RATIONAL MORAL ENLISTMENT

White suggests a rival approach to moral education to the one that Hand prefers. Whereas Hand is a rationalist, seeking reasons for moral commitment, White eschews reasons, preferring to 'ground' moral commitment in sentiment:

On this view, which I favour, caring for people's wellbeing is something we bring children up to feel positively about. Hand may, of course, agree with this as something we should do at a pre-rational stage of moral education, provided that good reasons are supplied later. Where I

depart from him is in removing the proviso. Coming to care is all the bedrock needed (White, 2016: 454 - 455).

*Contra* White, I do not think that moral education simply boils down to a matter of our bringing up children to care about other people's wellbeing if we ourselves happen to care about other people's wellbeing, but to an additional fact that we should care about one another's well-being. That makes our doing so a non-random fact, rather than simply the way it turned out in the roulette of values formation. We act as if things were valuable; we refrain from acting in ways that damage those things which we take to be valuable, and we act in ways which advance and preserve those things which we think are most valuable. We act as if beholden to obligations that we perceive, and refrain from acting in ways that seem to us impermissible; we act in spite of what we might prefer to do, for instance. It seems, furthermore, that we cannot jettison our evaluations entirely, or take ourselves to always merely 'project' value rather than recognize it. While it might be *psychologically* hard or even impossible to switch all of our values by caprice, I want to suggest that it is not a *logical* possibility to think that it would be *all right* to simply change our judgements randomly (in a hypothetical, random attitudinal-amendment machine, say), or to undo all of our judgements and have none. This is because, to think that, would already mean that they were not really our judgements *at all*. To make sense of a sincere civil rights advocate being willing to undergo the risk of becoming a racist, say, there would need to be some other sincerely held value that would be well served by entering the machine. Similarly, to think that all of our judgements merely reflect arbitrary preferences would be to have already abandoned them. It seems that we cannot step outside of our values. In having values, we are committed to our value judgements being right or wrong irrespective of what we happen to think, otherwise we would already be judging the objects that we value to be worthless. That said, we will likely find that our values can be made to form a more coherent whole than they do, by amending them in light of one another. Indeed, this process of mutual amendment (or reflective equilibrium) seems to be the best method for developing one's ethical view. Furthermore, it is a procedure that can be prosecuted by teachers and parents in moral education.

Also *contra* White, we should worry about dispensing with Hand's epistemic criterion in moral education. If it were dispensed with across the board we lose our most powerful means of objecting to evolution versus creationism being taught in a non-directive way, and to religious beliefs being taught in a directive way (both of which White agrees to). While I accept a broadly Handian framework within which the problem of non-rational moral enlistment can be framed, I do not find Hand's answer satisfying. I suggest that we need not have identified any moral foundation, and that students need not be presented with one in order for teachers to conduct rational moral enlistment within what I take to be a broadly Handian framework. I will propose in the next section that the moral intuitions that people have already do not require grounding in a self-evident moral truth, they are how things seem to people, and while they are defeasible seemings, they do serve as evidence that things are thus and so. Such intuitions need only to cohere together, and can potentially defeat one another. On this understanding, there is scope for moral enlistment on the strength of reasons, but without needing to identify a foundational moral truth at the outset. Coming across such a truth in the process is not precluded, although Rawls himself argued that it was exceedingly unlikely (Rawls, 1974).

## 5. REFLECTIVE EQUILIBRIUM AND JUSTIFICATION

In this section, I describe the procedure of reflective equilibrium, drawing principally on a four step account developed by Ben Spiecker and Jan Steutel (2001). I explain the various points of departure that I have from their rendering of the procedure, before explaining something about its attractions. Spiecker and Steutel present reflective equilibrium as a method we may use 'to find out whether ethical claims are justified or credible' (2001: 31).<sup>2</sup> They offer a four step account of the procedure, which I want to endorse in some respects, and revise in others:

The method of reflective equilibrium contains four basic steps: (1) the collection of initial ethical judgements; (2) filtering our initial judgements to arrive at considered ethical judgements; (3) construing the best fit between our considered judgements and ethical principles (narrow reflective equilibrium); and (4) figuring out what the best fit is between our

considered judgements, principles and background theories (wide reflective equilibrium) (2001: 36)

I will explain each of these steps together with my own amended version.

(1) First, Spiecker and Steutel require that when we ‘articulate and collect our initial ethical judgements,’ we should retain only those which ‘have a degree of credibility independently of their inferential relations with other beliefs’ (32). They call these our ‘initially tenable beliefs’. The idea of non-inferentially based degrees of plausibility is suspicious – outside of self-contradiction and self-consistency it is not clear what there is to choose between beliefs considered in isolation. Surely ideas could only be considered more or less tenable by satisfying some criteria that we would also have to be committed to, criteria that would also be among the material admitted among our initial commitments which can only derive further credibility from the procedure they are about to be submitted to. Developing this point, Spiecker and Steutel agree with Rawls that we should aim to bring into equilibrium only ‘those judgements in which our moral capacities are most likely to be displayed without distortion’ (Rawls, 1972: 47). The trouble with this idea is that it requires an account of our moral capacities and their proper functioning i.e. an account of the conditions under which our moral capacities generate the correct results. Such an account would presuppose the very things we are looking for: namely reliable ethical judgements. Instead, I submit that all of our initial judgements are equally credible until the procedure of reflective equilibrium has gone to work on them. While one may admit that some judgements are objectively mistaken, viewed from the internal perspective, they are still on a par with all other judgements. It will only be when we have our schema of correct judgements to pair off with our faculties and exercises of judgement that we will see what material conditions and cognitive equipment leads judgement awry. Then we may consider whether we ourselves might be defective in some respect in forming moral judgements and so would do best to defer to those better equipped than ourselves. That said some are very likely to be relevant:

Since the rightness or wrongness of an action, or the justice or injustice of an institution, depends on facts about it, judgments made in ignorance of these facts are unreliable.

Factors such as self-interest and emotional distress also make a judgment unreliable, because these factors can interfere with a person's assessment of the morally relevant considerations. (Scanlon, 2002: 145)

Following Rawls in his 1974 Presidential Address, 'The Independence of Moral Theory,' it is principles that 'match people's considered judgments and general convictions in reflective equilibrium,' that we are to seek (i.e. their judgments at all levels of generality). (Rawls, 1999: 289). A shift of his original statement in (Rawls, 1951)

(2) The second step seems proper, but for reasons other than Spiecker and Steutel give. They recommend we further cull our set of 'initially tenable beliefs,' to those 'in which we have the greatest confidence,' (2001: 32) which are 'our firmest convictions'. Distinguishing between *justified confidence* and *psychological confidence*, we must again ask: what *justifies* our confidence *apart* from the process of reflective equilibrium? Without an answer to this, we must again content ourselves to start out on our reflective process proper without having culled our set of initial beliefs. On the other hand, we should allow that *psychological confidence* is highly relevant to which beliefs we should attempt to bring into equilibrium. This is because it is those beliefs of which we are most confident that are most part of us. It is no good bringing into equilibrium views to which I am uncommitted, for then I would be uncommitted to the picture which emerges. If asked which judgements are really ours, we must admit to some complications. We may distinguish between the values implicit in our actions, those which we knowingly act on, and finally those that we want to act on.<sup>3</sup>

(3) The third step is where the real business of the procedure, as I understand it, gets underway: it is here that 'we try to construct an ethical theory by explicating ethical principles that account for our considered judgements' (34); this step 'may imply adjusting both the suggested ethical principles and our considered ethical judgements' (34); 'our ethical intuitions may prove to be irregular, not fully consistent and sometimes distorted, because of which they may lose their credibility' (35); 'eventually they fit together into one coherent ethical view' (35); 'The result is a reflective equilibrium between ethical intuitions and ethical principles: our intuitions do support the principles and the principles do increase the initial tenability of our intuitions' (35). Put differently, one already has a range of ethical

intuitions ranging in generality, for instance: Oswald should not have shot JFK; the Police Sniper should have shot the armed man on a shooting spree; murderers should be executed; drawn out, torturous executions are wrong. With these in hand, one then attempts to see which principles can be formulated to explain the most, and revise fewest, of these intuitions. Preference would be given to the simplest, least extensive and most explanatorily powerful set of principles. 'A moral theory has *power* when it yields judgments not included in the original data base' (Kagan, 1989: 11). Where one has intuitions which are left by the wayside, one would ideally have a theory of error; of how one came to have the mistaken view, and some respect of its generation in which it differs from the others in virtue of which it is not to be trusted, while they are (after all, it was not so long ago that they seemed to be on the same footing). In this way, a theory is generated from intuitions. This, however, is only 'narrow' reflective equilibrium.

(4) The fourth step that Spiecker and Steutel advocate, in which our intuitions are to be brought into 'wide' reflective equilibrium, is one that I embrace in part, and in part express reservation about. 'According to the criterion of comprehensiveness, also all kinds of other beliefs that have some inferential bearing on our considered judgements and ethical principles should be taken into account' (35). Spiecker and Steutel call these: *background theories*, which may be ethical or non-ethical. Apparently these can 'force us to break open our narrow reflective equilibrium and to adopt and revise different parts of it, including some considered convictions' (2001: 35). Why is this so? In respect of background moral theories, they suggest that 'the principle of mutual consent can be justified in terms of a Kantian ethical theory of the person, particularly the idea that persons have intrinsic dignity because they are rational agents' (35). 'This ethical background theory supports the principle of mutual consent on grounds to some degree independent of its match with the relevant judgements' (35). While this might be so, taking this to be the case means that one has a *hybrid theory* of moral justification in which reflective equilibrium does *not* do all of the justificatory work. One would need to be persuaded of Kant independently, and then if Kantian theory doesn't match one's intuitions, one will have to choose whether to jettison one's pre-theoretical intuitions as mere prejudice to be corrected by independent Kantian considerations. Unless one is persuaded by some Kantian theory, such a theory is irrelevant to

ethical justifications. It is the process of reflective equilibrium is all that is needed to a) generate theory and b) lend it what plausibility it has.

Spiecker and Steutel are more convincing on the count of non-moral background theories having relevance to our ethical views: 'justifying an ethical belief does not exclusively consists [sic] of showing it to cohere well with other ethical beliefs, but also essentially appealing to non-ethical views, especially to psychological, political and economic theories' (35). Take the following example from Stephen Law: 'The dispute over whether or not women should be allowed to vote was, in part, a dispute over whether women have the necessary intellectual skills to exercise that right properly' (Law, 2006: 118). They add the sensible requirement of a feasibility test: this is 'to make sure that living according to the ethical views we are defending is possible for creatures like us' (Spiecker and Steutel, 2001: 36).

What we have in reflective equilibrium is 'a coherence account of justification (as contrasted with an account of truth)' (Daniels). For Spiecker and Steutel: 'The more comprehensive and coherent an ethical conception is, and the better it takes into account our initially tenable beliefs, the stronger is the justification of its constituent parts' (33). Coherence 'is not just a function of logical consistency but also a matter of connectedness by inferential relations' they constitute 'a network of mutually supporting claims,' rather than a motley assortment of 'mutually indifferent judgements' (32), what Kagan calls a 'laundry list' (Kagan, 1989: 11); 'Other things being equal, a more comprehensive coherent set of beliefs is more credible than a less comprehensive one' (Spiecker and Steutel, 33). 'The credibility of our ethical view on human interaction would be increased if it could be integrated with those other [coherent sets of] views [on different topics] into a more comprehensive coherent belief set' (33). Concordantly, Gary Gutting has suggested that someone's ability to sustain a point of view at length is evidence in its favour:

The mere fact of developing a claim in some detail may serve to persuade us of its plausibility. Since more detail is likely to lead to problems ... the more thoroughly and extensively a claim is developed without encountering problems, the more likely it is to be correct. (Gutting, 2009: 77).

*Contra* the views of Coherentists about truth, it seems that coherence on a topic is not enough to render an account of things true, it is rather a precondition (even if, as Gutting suggests, an account of things gains credibility with detailed development). Coherentists hold that coherence is necessary and sufficient to render an account of things true. However, Coherentist's thesis seems to be incoherent and thus self-defeating as an account of truth: It is perfectly coherent to maintain that an account is true (only) by virtue of its correspondence to reality (which would render that account of truth correct on the Coherentist's analysis) but it is incoherent to maintain that truth is both a function (only) of correspondence to reality and (only) of its coherence. In order to defeat this style of objection, what the Coherentist about truth has to deny is the view that any other account of truth is even coherent. But that is an improbably ambitious task.

## 7. REFLECTIVE EQUILIBRIUM AND THE PROBLEM OF RATIONAL MORAL ENLISTMENT

How can the process of reflective equilibrium help contribute to the Problem of Rational Moral Enlistment? Compelling arguments making use of reflective equilibrium can be formulated and put to children in educational contexts, in fact they already are. Consider the following example after a not uncommon incident in which one child kicks another out of anger at having been pushed in front of in the queue for lunch. A teacher might ask: Would it be alright if things had been the other way around and so and so had kicked you? The answer might reasonably be: no, that would not be alright. The next question might be: why is it alright for you to kick so and so? The student then has to choose between it being acceptable for them to be kicked, it not being ok for them to kick so and so, or there being some material difference between the cases. Here the thought does not run on empathy, on imagining what it is like for others to undergo suffering and to feel for them, but to regard their wellbeing as in no wise less important than their own wellbeing on the grounds failure to produce any relevant difference between the cases. General explanatory principles can be formulated, such as: 'do not kick people', or more generally still 'do not harm'. Exceptions might be thought of to the rule, and it might then be modified in light of them.



This line of conversation, which may already seem quite familiar from recollections of having one's disputes with peers settled in the playground, can be systematized and followed more rigorously. The step in the process can be made explicit, and the task completed by comprehensively and critically. On this model, having informed themselves about an issue, students will (1) articulate their judgements about a range of cases with different kinds of features. They will be encouraged to set aside those of their judgments that were generated in emotional distress, in ignorance of facts, or which were motivated by self-interest. (2) Of these judgements they will again be encouraged to set aside those which they have the least confidence in, since they would not seem to bare out what the student really thinks. (3) Narrow reflective equilibrium would be sought between the pupil's judgements and simple, explanatorily powerful principles that would explain the most and revise the fewest of these judgements. (4) Finally, wide reflective equilibrium would be sought between students' considered judgements, explanatory principles and other non-ethical background theories which might render them more or less implausible. By following this procedure, children can be brought to have epistemically justified ethical views, thereby solving the Problem of Rational Moral Enlistment. In order to further defend this account I will, in Section 8, anticipate some objections to my argument.

## 8. OBJECTIONS TO REFLECTIVE EQUILIBRIUM AND REPLIES TO THEM

### Coherence ain't no justification

Richard Brandt (1979) and Richard Hare (1973) have observed that there must be something good or true on the whole about our initial intuitions, for the process not to be an exercise in systematic fiction writing, but worried that we have no reason to think that there is. Rawls himself requires that our 'considered judgements,' which are to bring into reflective equilibrium, be generated in circumstances favourable to judgment as such, granting them a prima face credibility; i.e. where we are informed about the facts of the matter, are in a position to deliberate on them uninterrupted, or pressured, and are disinterested (Rawls, 1971: 47-8). This is all as much as to say, we have no reason to doubt them. Indeed, I think Brandt and Hare put the burden of proof in the wrong direction: we have reason to accept that things are as they seem to us to be, all the more so to the extent that they can be rendered into a

coherent whole. When we consider cases like Auschwitz and ask whether we think that there is anything wrong with them really, or whether we *just happen* to have a con attitude rather than a pro attitude, that it might have turned out differently and ultimately not have mattered which attitude we took (as might be thought to be so with norms of etiquette), where we fear we would be doing something calamitous to unburden ourselves of our moral evaluations, we need not have a story about how our intuitions are formed, at least not ahead of time, it is enough that we regard them as having some force already, and then seek to see whether they pan out systematically. The next step is to see what in our nature and in our environment can systematically yield such judgements in a vindicatory fashion, as well as what might count as defects in our judgements when we systematically deviate from that pattern of judgement. But, 'it is premature to ask for such a story in ethics,' for 'we can provide no analogous causal story for credible judgments we make in other areas, including mathematics or logic' (Daniels, 2016). It is surprising that Utilitarians should take up these wholesale rejections of moral seemings since they rely on them much as anyone else. Indeed, in attacking them, the Utilitarian seems to deprive herself of any resources for building a theory at all. Consider Peter Singer's appeal to our intuitions with his child in the pond example, and the way he builds on this intuition in his paper, 'Famine and Affluence'. Indeed, there is no necessary tension between Utilitarianism and reflective equilibrium; reflective equilibrium might vindicate Utilitarianism. Some might hope to cast shade on reflective equilibrium by remarking that someone's astrological judgements cannot be rendered plausible through generating explanatory principles. I suggest that such judgements simply would not survive reflective equilibrium: one would reason oneself out of astrology if one were to think seriously about one's astrological judgements, attempting to bring them into wide reflective equilibrium. The process of reflective equilibrium will likely jettison many an intellectual seeming. As David Lewis has emphasized, one starts off with one's *own* judgements (indeed, where else could one start off?), and revises from there.

a theory cannot earn its credence just by its unity and economy. What credence it cannot earn, it must inherit. It is far beyond our powers to weave a brand new fabric

of adequate theory ex nihilo, so we must perforce conserve the one we've got (Lewis, 1986: 134)

One takes initially plausible, but not sacrosanct beliefs, and attempts to systematize them, revising or rejecting them if they resist such integration.<sup>4</sup>

### Conservativism

T.M. Scanlon observes that 'The charge of conservatism is based on the plausible idea that an adequate method for deciding what to believe about a subject must provide some standard with reference to which the current beliefs we happen to have can be judged and perhaps found wanting' (Scanlon, 2002: 150). On this view reflective equilibrium can look worryingly conservative, and incapable of taking its original judgements to be in error. However, the following considerations tell against that conclusion.

But first, our original judgements are not fixed, 'but are open to constant revision' (150), second 'the class of considered judgements includes judgements of all levels of generality' (150), and third we are to weigh them against them 'other plausible conceptions' and will have had a chance to 'assess the supporting ground' of these (150). As Scanlon concludes, 'It is difficult to imagine what source of criticism or justification is envisaged that the method of reflective equilibrium, so understood, would exclude' (150). However, it is also important to acknowledge that even small changes to a dataset can, once integrated consistently throughout, can yield massive changes to the data set as a whole.

### Relativism

Another common, but misguided objection to reflective equilibrium is that it courts moral relativism. Scanlon poses the apparent problem in this way:

Suppose I have carried through the process Rawls describes and found principles that are in reflective equilibrium with my considered judgments. It could still be claimed, not implausibly, that some other person who was equally well-informed might carry through the process just as conscientiously and reach a different result. Is the defender of reflective equilibrium not then

committed to the claim that the incompatible sets of principles we have reached are both justified? (152)

However, as Scanlon equally points out, ‘Accepting the method of reflective equilibrium does not commit me to the view that the principles this other person has reached are justified unless I judge that not only that person’s starting points but also all of the steps he or she made along the way are sound.’ (153). So the charge of relativism holds no water. It may be that people start off with different intuitions, and that the process will just bring people to different systems in light of that; this is a legitimate worry that I cannot completely diffuse. Indeed, while we might attempt to persuade one another, there is no guarantee that we will all be able to converge in our opinions through rational discourse, but relativism is no consequence of this.<sup>5</sup>

It may be that neither of us is making any mistake of method. We may each be bringing our opinions into reflective equilibrium in the most careful possible way, taking account of all the arguments, distinctions, and counterexamples. But one of us, at least, is making a mistake of fact (Lewis, 1983: xi)

#### Is reflective equilibrium really directive?

Teachers cannot be sure ahead of time that they can use their students’ existing judgements to support any particular principles, or that those principles can in turn be used to revise a certain set of particular judgements. But certainly the procedure of reflective equilibrium can be done better or worse, it is not a matter of taste or caprice that a certain array of judgments fit best with a certain simple set of principles. It is in guiding this process and in showing the student connections, in leading them to coherence, simplicity and power, within their judgements, that the teacher does something that is well described as directive, even if that does not mean imparting their own set of moral judgements. But our question was whether we could resolve the Problem of Rational moral recruitment, and it seems we can do just this.

## **9. CONCLUSION**

As Hand argues, there is indeed a problem regarding how one can educate children to recognize the requirements of morality without resorting only to non-rational means of persuasion. This is because, firstly, directive teaching ought to be by means of rational persuasion, such that, ultimately, the reasons for the belief can be understood by the pupils, and secondly, what morality requires and why seems to be a matter of reasonable dispute. Hand proposed to solve the problem by arguing that *some* moral standards are beyond reasonable dispute, claiming to find their foundational rational justification in the amelioration of ‘breakdowns in cooperation and outbreaks of conflict’ within social groups (2014b: 528, 530). However, the moral foundation that Hand claims to find in ameliorating the problem of sociality is 1) less persuasive than many of the prohibitions, obligations and other standards that it is meant to ground, 2) fails to ground intuitions that are more appealing than this way of grounding them, and 3) conflicts with other intuitions that are more appealing than this way of grounding them. For these reasons, it cannot be the indefeasible, self-evident basis for morality that Hand claims it to be. While we have good reason to accept Hand’s formulation of the Problem of Rational Moral Enlistment, and good reasons to reject his solution to it, we are not thereby at a loss. As I have argued, reflective equilibrium represents a promising way forward. While I do not argue that moral foundationalism is impossible, I contend that Hand has not found a foundational reason, and furthermore that we do not need one. Whereas Hand hopes to ground rational moral enlistment in a single, self-evident foundational justification of some moral standards, I instead hope to facilitate rational moral enlistment through a rational procedure which starts with students’ existing moral commitments and attempts to revise or expand them through a certain kind of critical reflection.

There is an ambiguity of scope concerning reflective equilibrium. On the one hand, teachers could attempt to effect reflective equilibrium in each pupil, on the other hand the teacher could attempt to bring their own views into reflective equilibrium and impart those, communities of local, national or international size could attempt to bring their views into reflective equilibrium and employ teachers to impart those. It seems to me that in attempting to persuade children of the advantages of their own views, teachers will have to either have to give reasons that connect with students’ own belief sets, or they will have to suggest that students’ beliefs sets need not be appealed to, perhaps because the

community's having brought its views into reflective equilibrium provides sufficient epistemic warrant to impart the views. However, in the absence of an error theory which identifies the defects in children's moral belief formation, there is no particular reason to suspect that the children's beliefs are likely to be more at fault than the community's: it is exactly by appealing to and bringing into equilibrium how things seem to the children that rational moral enlistment is to be effected. As their views achieve equilibrium, they begin to have rational warrant for holding them.

## REFERENCES

- Brandt, R. (1979) *A Theory of the Good and the Right*, Oxford: Oxford University Press.
- Daniels, Norman, 'Reflective Equilibrium', in Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*, Winter 2016 Edition, URL = <https://plato.stanford.edu/archives/win2016/entries/reflective-equilibrium/>.
- Frankfurt, H.G. (1988) 'Freedom of the will concept of a person', in *The Importance of What We Care About: Philosophical Essays*, Cambridge: Cambridge University Press.
- Gutting, G. (2009) *What Philosophers Know: Case Studies in Recent Analytic Philosophy*, Cambridge: Cambridge University Press.
- Hand, M. (2003) 'A Philosophical objection to faith schools'. *Theory and Research in Education* 1: 89–99.
- Hand, M. (2014a) 'Afterwords: Response to Warnick and Smith'. *Educational Theory* 64: 425–426.
- Hand, M. (2014b) 'Towards a theory of moral education'. *Journal of Philosophy of Education* 48: 519–532.
- Hare, R.M. (1973) 'Rawls' Theory of Justice'. *Philosophical Quarterly* 23: 144–55; 241–51.
- Kagan, S. (1989) *The Limits of Morality*, Oxford: Clarendon Press.
- Law, S. (2006) *The War for Children's Minds*, Abingdon: Routledge.
- Lewis, D. (1983) *Philosophical Papers, Volume One*, Oxford: Oxford University Press.
- Lewis, D. (1986) *On the Plurality of Worlds*, Oxford: Blackwell Publishers.
- Lowe, E. J. (2002) *A Survey of Metaphysics*, Oxford: Oxford University Press.

- McMahan, J. (2000) 'Moral intuition'. in H. LaFollette (ed.), *Blackwell Guide to Ethical Theory*, Oxford: Blackwell.
- Muir, J. (2016) 'Islamic State group: The full story', published: 20 June 2016 Available at: <http://www.bbc.co.uk/news/world-middle-east-35695648> [Last accessed 24/05/2017]
- Nagel, T. (1986) *The View from Nowhere*, Oxford: Oxford University Press.
- Nolan, D. (2005) *David Lewis*, Cheshire: Acumen Publishing.
- Rawls, J. (1951) 'Outline of a decision procedure for ethics', *Philosophical Review* 60: 177–97, reprinted in *Collected Papers*, 1999, pp. 1–19.
- Rawls, J. (1972) *A Theory of Justice*, Oxford: Clarendon Press.
- Rawls, J. (1971) *A Theory of Justice*, 2<sup>nd</sup> Edition 1999, Cambridge, MA: Harvard University Press.
- Rawls, J. (1974) 'The Independence of Moral Theory', *Proceedings and Addresses of the American Philosophical Association*, 47: 5–22, in *Collected Papers*, 1999, pp. 286–302
- Rawls, J. (1999) *Collected Papers*, ed. Samuel Freeman, Cambridge, MA: Harvard University Press.
- Scanlon, T.M. (2002) 'Rawls on justification', in Freeman, S. (ed.) *The Cambridge Companion to Rawls*, Cambridge: Cambridge University Press: pp. 139–167.
- Spiecker, B. and J. Steutel, (2001) 'Reflective equilibrium as a method of philosophy of education', in F. Heyting, D. Lenzen, and J. White (eds) *Methods in Philosophy of Education*, London: Routledge.
- Taylor, R. (2017) 'Indoctrination and social context: A system-based approach to identifying the threat of indoctrination and the responsibilities of educators'. *Journal of Philosophy of Education* 51: 38–58.
- White, J. (2016) 'Moral education and education in altruism: Two replies to Michael Hand'. *Journal of Philosophy of Education* 50: 448–460.

#### BIOGRAPHICAL NOTES

Dr John Tillson is a Teaching Fellow in the Department of Philosophy at the University of Warwick. He has taught philosophy and philosophy of education at Dublin City University where, in 2015, he completed his PhD thesis, *Children, Religion and the Ethics of Influence*, on a scholarship from the Irish Centre for Religious Education. He has published peer reviewed articles on teaching controversial issues, Curriculum Theory, Ethics Education, Religious Education, and on Bernard Williams' metaphysics. He has also presented papers at academic proceedings around Europe, North America and Australasia. He has been awarded honorary lifetime membership of the British Humanist Association, the Conway Hall Ethical Society, and of Trinity College Dublin's Metaphysical Society.

FUNDING STATEMENT: No funding was received in connection with developing this paper.

---

<sup>1</sup> Those too can be formed without reasons and arguments, but because they are not beliefs, and do not aim at truth, Hand thinks that they are immune to indoctrination. He might call these practises ‘conditioning,’ and regard our being non-rationally initiated into them as morally wrong only where those habits and dispositions themselves were morally wrong. The distinction might seem to give scope to Hand to allow room for moral education without indoctrination: teachers might simply condition their students into respecting the demands of morality. However, those kinds of formation which encourage closed-mindedness are well regarded as miseducational, and are at least plausible candidates for the label ‘indoctrinatory’ (Taylor, 2017). Since habits or dispositions can be cultivated in ways that close people off from critically reflecting on them, so too can they be indoctrinated, and so it seems that the distinction in no wise salvages non-rational moral enlistment. At any rate, Hand regards any moral education is incomplete which fails to teach for cognitive assent to moral judgements.

<sup>2</sup> To be sure, reflective equilibrium is concerned with what can justify judgements, not about what constitutes reality. We will come back to this point in connection with David Lewis’ use of reflective equilibrium in his meta-metaphysics. Roughly, we allow that we might have all the justification we could ever hope for, but still not have the truth.

<sup>3</sup> On this point, compare Harry Frankfurt’s discussion of first and second order desires in ‘Free Will and the Concept of a Person’ (1988). Frankfurt identifies a person most closely with the set of wants they want to have, but that move is contestable – arguably a person ought to be most closely identified with the sorts of consideration which actually motivate their actions.

<sup>4</sup> For an interesting discussion of how far this enterprise can and should be undertaken, see Thomas Nagel’s *The View from Nowhere* (1986) in which he suggests that false reductions, and denials of the most evident features of the world can result from pursuing it over-zealously.

<sup>5</sup> For a reconstruction and illuminating discussion of David Lewis’ general philosophical method, see Chapter 9 of *David Lewis* (Nolan, 2005).